Comparison of Search Ability between Genetic Fuzzy Rule Selection and Fuzzy Genetics-Based Machine Learning

Yusuke Nojima, *Member, IEEE*, Hisao Ishibuchi, *Member, IEEE* and Isao Kuwajima, *Student Member, IEEE*

***Abstract*— We developed two GA-based schemes for the design of fuzzy rule-based classification systems. One is genetic rule selection and the other is genetics-based machine learning (GBML). In our genetic rule selection scheme, first a large number of promising fuzzy rules are extracted from numerical data in a heuristic manner as candidate rules. Then a genetic algorithm is used to select a small number of fuzzy rules. A rule set is represented by a binary string whose length is equal to the number of candidate rules. On the other hand, a fuzzy rule is denoted by its antecedent fuzzy sets as an integer substring in our GBML scheme. A rule set is represented by a concatenated integer string. In this paper, we compare these two schemes in terms of their search ability to efficiently find compact fuzzy rule-based classification systems with high accuracy. The main difference between these two schemes is that GBML has a huge search space consisting of all combinations of possible fuzzy rules while genetic rule selection has a much smaller search space with only candidate rules.**

I. Introduction

Since some pioneering studies [1]-[5] in the early 1990s, genetic algorithms [6], [7] have been successfully used in the design of fuzzy rule-based systems [8]-[12]. Those GA-based approaches are often referred to as genetic fuzzy systems [10]. For the design of compact fuzzy rule-based classification systems with high accuracy, we proposed two schemes: One is genetic rule selection [13]-[16] and the other is genetics-based machine learning (GBML) [17]-[21].

In genetic rule selection, first a large number of fuzzy rules are extracted from numerical data in a heuristic manner as candidate rules. Then a small number of fuzzy rules are selected by a genetic algorithm. Genetic rule selection for classification problems was first formulated as a single-objective optimization problem with the following weighted sum fitness function in [13], [14]:

$$f(S) = w_1 \cdot f_1(S) - w_2 \cdot f_2(S), \qquad (1)$$

where S is a subset of candidate rules, $f_1(S)$ is the number of correctly classified training patterns by S, $f_2(S)$ is the number of selected fuzzy rules in S, and w_1 and w_2 are prespecified positive weights. A single-objective genetic algorithm was used to find the optimal rule set of the fuzzy rule selection problem in (1). The single-objective problem in (1) was generalized as two-objective rule selection in [15] where a two-objective genetic algorithm was used to find non-dominated rule sets with respect to $f_1(S)$ and $f_2(S)$. The two-objective formulation in [15] was further extended to three-objective rule selection [16] where the minimization of the total number of antecedent conditions (i.e., the total rule length) was introduced as the third objective $f_3(S)$.

Whereas candidate rules are heuristically generated in advance in genetic rule selection, GBML generates fuzzy rules from existing ones by genetic operations. For the design of compact fuzzy rule-based classification systems with high accuracy in the fuzzy GBML framework, we proposed a Michigan-style algorithm in [17], [18] where each fuzzy rule (which was denoted by its antecedent fuzzy sets as an integer string) was handled as an individual. We also proposed a Pittsburgh-style algorithm to optimize rule sets in [19]-[21] where each rule set (which was represented by an integer string concatenating several fuzzy rules) was handled as an individual. In our Pittsburgh-style algorithm, a Michigan-style algorithm was utilized as a kind of mutation. Thus our fuzzy GBML algorithm can be viewed as a hybrid version of the Michigan and Pittsburgh approaches.

In this paper, we compare the two GA-based schemes (i.e., genetic rule selection and GBML) with each other in terms of their search ability to efficiently find compact fuzzy rule-based classification systems with high accuracy. This paper is organized as follows. First we explain fuzzy rule-based classification systems in Section II. Next we explain genetic rule selection in Section III. Then we explain GBML in Section IV. These two GA-based schemes are compared with each other through computational experiments on some test problems in the UCI machine learning repository in Section V. Finally we conclude this paper in Section VI.

II. Fuzzy Rule-Based Classification Systems

A. Pattern Classification Problem

Let us assume that we have m training patterns $\mathbf{x}_p = (x_{p1}, x_{p2}, ..., x_{pn})$, $p = 1, 2, ..., m$ from M classes in an n-dimensional continuous pattern space where x_{pi} is the attribute value of the p-th training pattern for the i-th attribute. For the simplicity of explanation, we assume that all the attribute values have already been normalized into real numbers in the unit interval [0, 1].

B. Fuzzy Rules for Pattern Classification Problem

For our n-dimensional pattern classification problem, we use fuzzy rules of the following form [22]:

This work was partially supported by Grant-in-Aid for Scientific Research on Priority Areas: KAKENHI (18049065)

Yusuke Nojima, Hisao Ishibuchi and Isao Kuwajima are with the Department of Computer Science and Intelligent Systems, Graduate School of Engineering, Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan (nojima@cs.osakafu-u.ac.jp, hisaoi@cs.osakafu-u.ac.jp, kuwajima@ci.cs.osakafu-u.ac.jp).

0-7803-9719-3/06/$20.00 ©2006 IEEE

Rule R_q: If x_1 is A_{q1} and ... and x_n is A_{qn} then Class C_q with CF_q, (2)

where R_q is the label of the q-th fuzzy rule, $\mathbf{x} = (x_1, ..., x_n)$ is an n-dimensional pattern vector, A_{qi} is an antecedent fuzzy set, C_q is a class label, and CF_q is a rule weight (i.e., certainty grade). We also denote the fuzzy rule R_q in (2) as $\mathbf{A}_q \Rightarrow \text{Class } C_q$ where $\mathbf{A}_q = (A_{q1}, ..., A_{qn})$. The rule weight CF_q has a large effect on the accuracy of fuzzy rule-based classification systems as shown in [22]-[24]. For other types of fuzzy classification rules, see [22], [25], [26].

Since we usually have no *a priori* information about an appropriate granularity of the fuzzy discretization for each attribute, we simultaneously use multiple fuzzy partitions with different granularities as shown in Fig. 1. In addition to the 14 fuzzy sets in Fig. 1, we also use the domain interval [0, 1] itself as an antecedent fuzzy set in order to represent a *don't care* condition. As a result, we have the 15 antecedent fuzzy sets for each attribute.

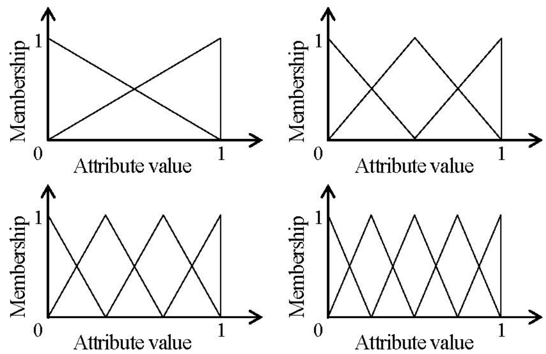


Fig. 1. Four fuzzy partitions used in our computational experiments.

C. Fuzzy Rule Generation

Since we have the 15 antecedent fuzzy sets for each attribute of our n-dimensional pattern classification problem, the total number of combinations of the antecedent fuzzy sets is 15^n. Each combination is used in the antecedent part of the fuzzy rule in (2). Thus the total number of possible fuzzy rules is also 15^n. The consequent class C_q and the rule weight CF_q of each fuzzy rule R_q are specified from the given training patterns in the following heuristic manner.

First we calculate the compatibility grade of each pattern $\mathbf{x}_p$ with the antecedent part $\mathbf{A}_q$ of the fuzzy rule R_q using the product operation as

$$\mu_{\mathbf{A}_q}(\mathbf{x}_p) = \mu_{A_{q1}}(x_{p1}) \cdot \ ... \ \cdot \mu_{A_{qn}}(x_{pn}), \quad (3)$$

where $\mu_{A_{qi}}(\cdot)$ is the membership function of A_{qi}.

Next the confidence of the fuzzy rule $\mathbf{A}_q \Rightarrow \text{Class } h$ is calculated for each class ($h = 1,2,...,M$) as follows [22]:

$$c(\mathbf{A}_q \Rightarrow \text{Class } h) = \frac{\sum_{\mathbf{x}_p \in \text{Class } h} \mu_{\mathbf{A}_q}(\mathbf{x}_p)}{\sum_{p=1}^{m} \mu_{\mathbf{A}_q}(\mathbf{x}_p)}. \quad (4)$$

The consequent class C_q is specified by identifying the class with the maximum confidence:

$$c(\mathbf{A}_q \Rightarrow \text{Class } C_q) = \max_{h=1,2,...,M} \{c(\mathbf{A}_q \Rightarrow \text{Class } h)\}. \quad (5)$$

The consequent class C_q can be viewed as the dominant class in the fuzzy subspace defined by the antecedent part $\mathbf{A}_q$. When there is no pattern in the fuzzy subspace defined by $\mathbf{A}_q$, we do not generate any fuzzy rules with $\mathbf{A}_q$ in the antecedent part. This specification method of the consequent class of fuzzy rules has been used in many studies since [27].

Different specifications of the rule weight CF_q have been proposed and examined in the literature. We use the following specification because good results were reported by this specification in the literature [22], [24]:

$$CF_q = c(\mathbf{A}_q \Rightarrow \text{Class } C_q) - \sum_{\substack{h=1 \\ h \neq C_q}}^{M} c(\mathbf{A}_q \Rightarrow \text{Class } h). \quad (6)$$

D. Classification of New Patterns

Let S be a set of fuzzy rules of the form in (2). When a new pattern $\mathbf{x}_p$ is presented to S, $\mathbf{x}_p$ is classified by a single winner rule R_w, which is chosen from S as follows:

$$\mu_{\mathbf{A}_w}(\mathbf{x}_p) \cdot CF_w = \max\{\mu_{\mathbf{A}_q}(\mathbf{x}_p) \cdot CF_q \mid R_q \in S\}. \quad (7)$$

When the rule set S includes no compatible fuzzy rule with $\mathbf{x}_p$, the classification of $\mathbf{x}_p$ is rejected. The classification of $\mathbf{x}_p$ is also rejected when multiple rules with different consequent classes have the same maximum value in (7).

III. Genetic Fuzzy Rule Selection

We use the following weighted sum fitness function in genetic rule selection and GBML to evaluate each rule set S:

$$f(S) = w_1 \cdot f_1(S) - w_2 \cdot f_2(S) - w_3 \cdot f_3(S), \quad (8)$$

where $f_1(S)$ is the number of correctly classified training patterns by S, $f_2(S)$ is the number of fuzzy rules in S, $f_3(S)$ is the total rule length of fuzzy rules in S, and w_1, w_2 and w_3 are positive weights. The rule length of each fuzzy rule means the number of its antecedent conditions excluding *don't care* conditions.

Using the fuzzy rule generation procedure in Subsection II.C, we can specify the consequent class and the rule weight for each of the 15^n combinations of the 15 antecedent fuzzy sets. The design of fuzzy rule-based classification systems can be viewed as finding the optimal subset of the 15^n fuzzy rules with respect to the fitness function in (8). Since any subset of the 15^n fuzzy rules is a feasible fuzzy rule-based classification system, the size of the search space is 2^{15^n}.

When n is large, this search space is intractably large. Thus we do not use all the 15^n fuzzy rules but only a prespecified number of promising fuzzy rules as candidate rules in genetic rule selection. When the interpretability of fuzzy rule-based classification systems is important, short rules with a few antecedent conditions are preferable to long rules with many conditions. Thus we use only short fuzzy rules of length $L_{\max}$ or less. In computational experiments, the value of $L_{\max}$ is specified as $L_{\max} = 3$ except for the case of the sonar data where $L_{\max} = 2$. We use these different specifications of $L_{\max}$ because the sonar data have much more attributes (i.e., 60 attributes) than the other data sets used in our computational experiments (e.g., 9 attributes in the glass data set and 13 attributes in the wine data set).

Among short fuzzy rules of length $L_{\max}$ or less, we only choose a prespecified number of candidate rules using a heuristic rule evaluation criterion. In this paper, we generate the best 300 fuzzy rules for each class with respect to the following heuristic rule evaluation criterion:

$$f(R_q) = s(\mathbf{A}_q \Rightarrow \text{Class } C_q) - \sum_{\substack{h=1 \\ h \neq C_q}}^{M} s(\mathbf{A}_q \Rightarrow \text{Class } h), \quad (9)$$

where $s(\mathbf{A}_q \Rightarrow \text{Class } h)$ is the support of the fuzzy rule $\mathbf{A}_q \Rightarrow \text{Class } h$, which is defined as follows [22]:

$$s(\mathbf{A}_q \Rightarrow \text{Class } h) = \frac{\sum_{\mathbf{x}_p \in \text{Class } h} \mu_{\mathbf{A}_q}(\mathbf{x}_p)}{m}. \quad (10)$$

The criterion in (9) can be viewed as a simplified version of a rule evaluation criterion used in an iterative fuzzy GBML algorithm called SLAVE [28]. Of course, we can use other criteria such as the support and the confidence as in data mining [29]-[31].

Using (9), we generate 300 fuzzy rules for each class (i.e., 300M fuzzy rules in total for an M-class problem) as candidate rules. Let S be a subset of N candidate rules where $N = 300M$. Any subset S can be denoted by a binary string of length N as $S = s_1 s_2 \cdots s_N$ where $s_j = 1$ and $s_j = 0$ mean the inclusion of the j-th candidate rule in S and its exclusion from S, respectively. Such a binary string is handled as an individual in genetic rule selection.

As we have already explained, each subset S is evaluated by the fitness function in (8). We use a single-objective genetic algorithm to search for the optimal subset of the N candidate rules with respect to (8). First an initial population is randomly generated. Then an offspring population is generated from the current population by binary tournament selection, uniform crossover and bit-flip mutation. We use biased mutation probabilities to efficiently decrease the number of fuzzy rules in each subset. More specifically, the mutation probability is specified as 0.1 for the mutation from 1 to 0, and 0.001 for that from 0 to 1. The $(\mu + \lambda)$-ES generation update mechanism is used to construct the next population. The population size is specified as $\mu = \lambda = 200$ in our computational experiments.

IV. Fuzzy Genetics-Based Machine Learning

Whereas the size of the search space is decreased from 2^{15^n} to 2^{300M} by the candidate rule prescreening phase in genetic rule selection, our fuzzy GBML algorithm has the original search space of size 2^{15^n}. In this section, we briefly explain our GBML algorithm. For details, see [20], [21].

Each fuzzy rule is represented by its antecedent fuzzy sets as an integer substring. Since the consequent part and the rule weight of each fuzzy rule are easily determined from the given training patterns by the heuristic rule generation procedure in Subsection II.C, they are not included in the substring. Thus the length of the substring is the same as the number of attributes (i.e., n). Each rule set is represented as a concatenated integer string where each substring of length n denotes a single fuzzy rule. The length of the concatenated string is nK when it includes K fuzzy rules.

We use the following fuzzy GBML algorithm in our computational experiments:

Step 1: Generation of an Initial Population

Each fuzzy rule in an initial rule set is generated from a randomly chosen training pattern in a heuristic manner [20]. An antecedent fuzzy set for each attribute is probabilistically chosen from the 15 antecedent fuzzy sets according to their compatibility grades with the attribute value of the chosen training pattern. The consequent class and the rule weight of each fuzzy rule are specified by the heuristic rule generation procedure in Subsection II.C. An initial population consists of 200 rule sets, each of which has 20 fuzzy rules.

Step 2: Evaluation of Each Rule Set

Each rule set in the current population is evaluated by the fitness function in (8).

Step 3: Genetic Operations

From the current population, 200 offspring rule sets are generated in the following manner:

3.1. Selection: A pair of parent rule sets are selected from the current population by binary tournament selection.

3.2. Crossover: An offspring rule set is generated by inheriting a randomly specified number of fuzzy rules from each parent. The offspring rule set can inherit up to 40 fuzzy rules from the parents. This means that the upper limit on the number of fuzzy rules in each rule set is specified as 40.

3.3. Mutation: Each antecedent fuzzy set in the offspring rule set is randomly replaced with another fuzzy set with a prespecified mutation probability. The mutation probability is specified as $1/n$ where n is the number of attributes.

3.4. Michigan Part: A single iteration of a Michigan-style fuzzy GBML algorithm is applied to the offspring rule set with the probability of 0.5. For details, see [20], [21].

Step 4: Generation Update

The next population is constructed by choosing the best 200 rule sets from the current population and the offspring population. That is, the same $(\mu + \lambda)$-ES generation update mechanism as in the case of genetic rule selection is used with $\mu = \lambda = 200$. When the prespecified stopping condition is not satisfied, return to Step 2.

V. Computational Experiments

A. Data Sets

We use six data sets in Table I: Wisconsin breast cancer (Breast W), diabetes (Diabetes), glass identification (Glass), Cleveland heart disease (Heart C), sonar (Sonar), and wine recognition (Wine). These six data sets are available from the UCI machine learning repository. Data sets with missing values are marked by "*" and "**" in the third column of Table I. We normalize all attribute values into real numbers in the unit interval [0, 1] in our computational experiments of this paper.

TABLE I
DATA SETS USED IN OUR COMPUTATIONAL EXPERIMENTS.

Data set	Attributes	Patterns	Classes
Breast W	9	683*	2
Diabetes	8	768**	2
Glass	9	214	6
Heart C	13	297*	5
Sonar	60	208	2
Wine	13	178	3

* Incomplete patterns with missing values are not included.
** Some suspicious patterns with an attribute value “0” are included.

B. *Conditions of Computational Experiments*

The two GA-based schemes are applied to each data set using the following common parameter specifications:

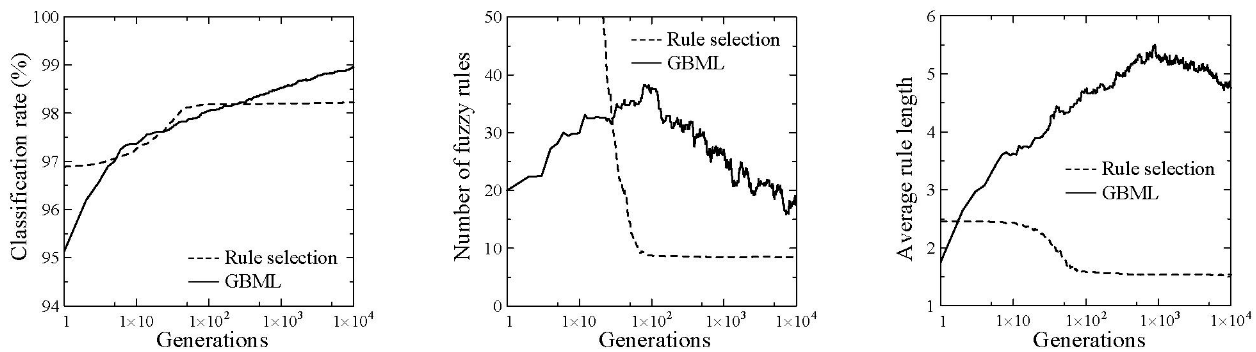
Population size: 200,
Crossover probability: 0.9,
Weight vector: $(w_1, w_2, w_3) = (100, 1, 1)$,
Stopping condition: 10000 generations.

In the execution of each GA-based scheme, 200×10000 rule sets are examined in their single run. All the given patterns in each data set are used as training patterns.

C. *Experimental Results*

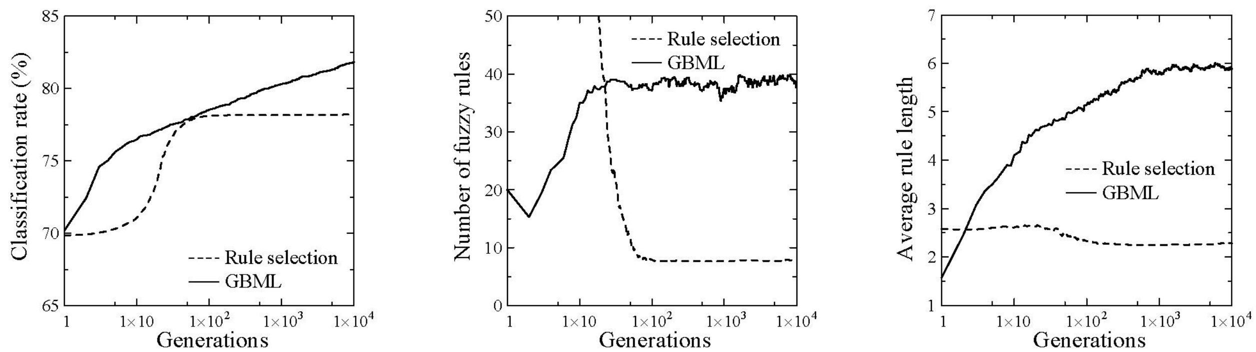
Experimental results are summarized in Figs. 2-7 where the average classification rate, the average number of fuzzy rules and the average rule length of the elite rule set in each generation over 20 independent runs are shown. In the left plots of all the six figures (i.e., Figs. 2-7 (a)), the dashed lines have similar shapes (i.e., S-shape learning curves). That is, the classification rates were rapidly improved by genetic rule selection during the first 100 generations. On the other hand, the classification rates of our GBML algorithm were gradually improved throughout 10000 generations (see the solid lines in Figs. 2-7 (a)). If we compare the two schemes at the 100th generation, genetic rule selection seems to be better than or comparable to GBML in Figs. 2-7 (a). At the 10000th generation, however, GBML seems to be better than or comparable to genetic rule selection in Figs. 2-7 (a).

In the center plots, we can observe a rapid decrease in the number of fuzzy rules by genetic rule selection during the first 100 generations (see the dashed lines in Figs. 2-7 (b)). Genetic rule selection found rule sets with less fuzzy rules than GBML in almost all cases (except for the final results in Fig. 7 (b)). From the right plots, we can see that genetic rule selection found shorter rules than GBML for all the six data sets (see Figs. 2-7 (c)). The average rule length was gradually increased by GBML until high classification rates were achieved (see the solid lines in Figs. 2-7 (c)). This contrasts with the results by genetic rule selection. From these observations, we can see that GBML found more accurate rule sets with higher complexity than genetic rule selection in the long run. Once GBML found accurate rule sets, it tried to decrease their complexity as shown in Fig. 7.



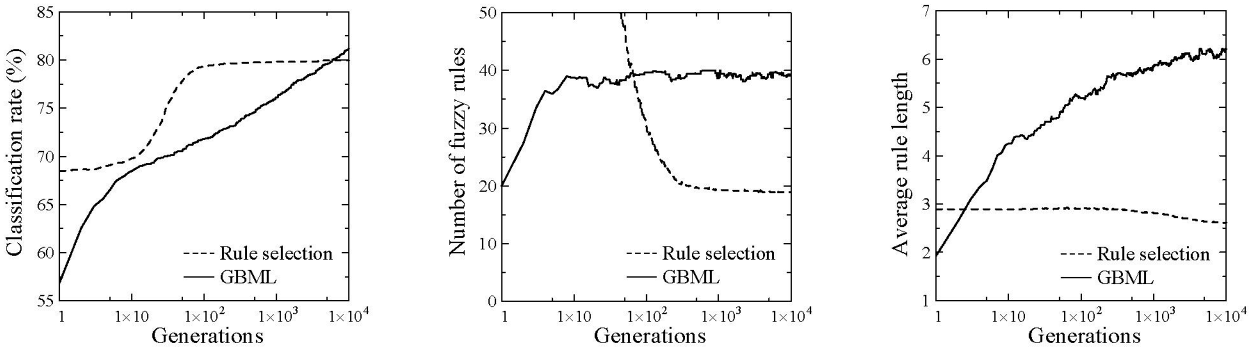
(a) Classification rate on training patterns. (b) Number of fuzzy rules. (c) Average rule length.

Fig. 2. Experimental results on the Wisconsin breast cancer data (Breast W).



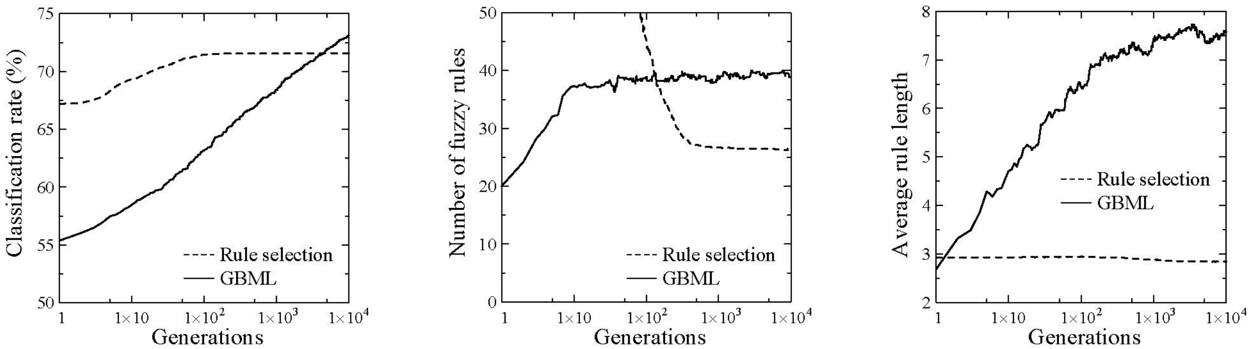
(a) Classification rate on training patterns. (b) Number of fuzzy rules. (c) Average rule length.

Fig. 3. Experimental results on the diabetes data (Diabetes).



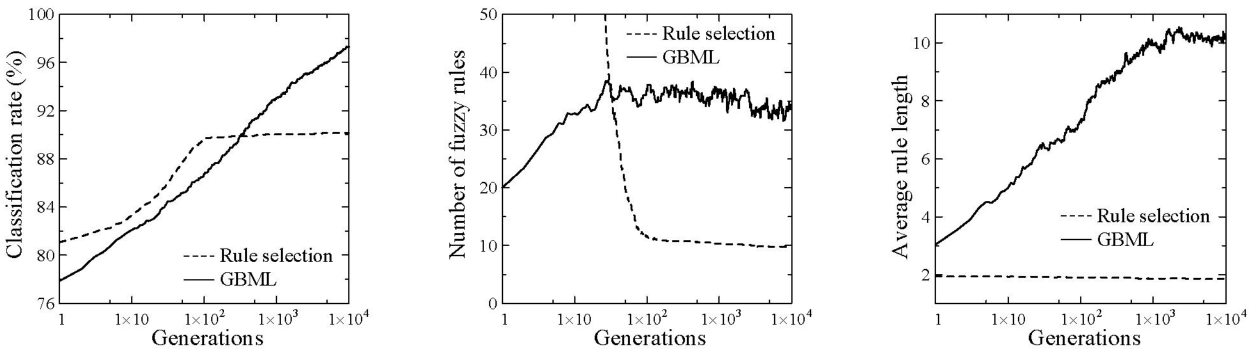
(a) Classification rate on training patterns. (b) Number of fuzzy rules. (c) Average rule length.

Fig. 4. Experimental results on the glass identification data (Glass).



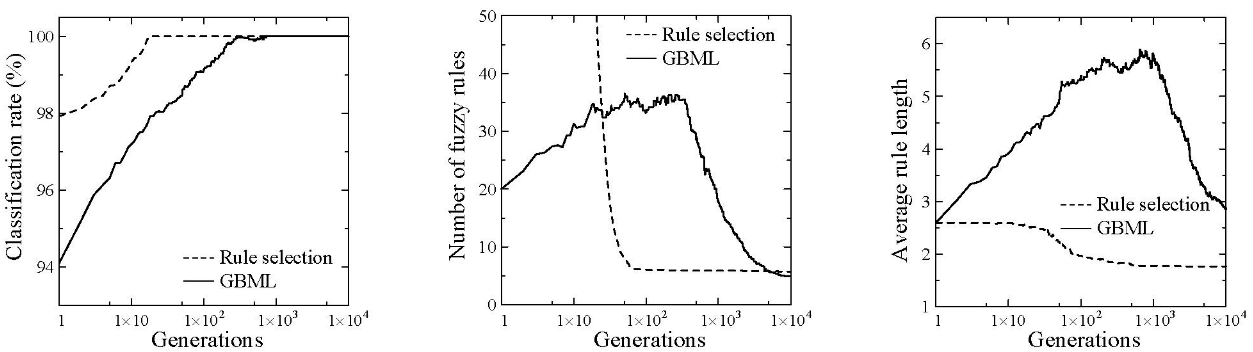
(a) Classification rate on training patterns. (b) Number of fuzzy rules. (c) Average rule length.

Fig. 5. Experimental results on the Cleveland heart disease data (Heart C).



(a) Classification rate on training patterns. (b) Number of fuzzy rules. (c) Average rule length.

Fig. 6. Experimental results on the sonar data (Sonar).



(a) Classification rate on training patterns. (b) Number of fuzzy rules. (c) Average rule length.

Fig. 7. Experimental results on the wine recognition data (Wine).

VI. Conclusions

We compared two GA-based schemes for the design of compact fuzzy rule-based classification systems with high accuracy. One is genetic rule selection and the other is genetics-based machine learning (GBML). Experimental results showed that each scheme has its own advantages and disadvantages. Genetic rule selection can efficiently find good rule sets within a small number of generations. In the long run, however, GBML comes to outperform genetic rule selection. These observations suggest an idea of hybridizing these two schemes in the following manner. First genetic rule selection is used to efficiently find good rule sets. Then GBML is executed to further improve the obtained rule sets by genetic rule selection.

Since the comparison between the two schemes was performed only in terms of their search ability, other important issues were not discussed in depth. One of such issues is the interpretability of obtained rule sets. As we briefly mentioned in Section V, genetic rule selection found a smaller number of shorter rules than GBML. That is, more interpretable rule sets were obtained by genetic rule selection. Whereas we always used homogeneous symmetric triangular membership functions, they are not necessarily an optimal choice. For example, we may need their adjustment if our goal is the accuracy maximization. Comparison with other genetic fuzzy systems is also a topic of future research.

References

[1] P. Thrift, "Fuzzy logic synthesis with genetic algorithms," *Proc. of 4th International Conf. on Genetic Algorithms*, pp. 509-513, University of California, San Diego, CA, July 13-16, 1991.

[2] C. L. Karr, "Design of an adaptive fuzzy logic controller using a genetic algorithm," *Proc. of 4th International Conf. on Genetic Algorithms*, pp. 450-457, University of California, San Diego, CA, July 13-16, 1991.

[3] M. Valenzuela-Rendon, "The fuzzy classifier system: A classifier system for continuously varying variables," *Proc. of 4th International Conf. on Genetic Algorithms*, pp. 346-353, University of California, San Diego, CA, July 13-16, 1991.

[4] D. S. Feldman, "Fuzzy network synthesis with genetic algorithms," *Proc. of 5th International Conf. on Genetic Algorithms*, pp. 312-317, University of Illinois, Urbana-Champain, IL, July 17-21, 1993.

[5] A. Parodi and P. Bonelli, "A new approach to fuzzy classifier systems," *Proc. of 5th International Conf. on Genetic Algorithms*, pp. 223-230, University of Illinois, Urbana-Champain, IL, July 17-21, 1993.

[6] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI, 1975.

[7] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989.

[8] F. Herrera and J. L. Verdegay (eds.), *Genetic Algorithm and Soft Computing*, Springer, Heidelberg, 1996.

[9] W. Pedrycz (ed.), *Fuzzy Evolutionary Computation*, Kluwer, Boston, 1997.

[10] O. Cordon, F. Herrera, F. Hoffman, and L. Magdalena, *Genetic Fuzzy Systems*, World Scientific, Singapore, 2001.

[11] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena (eds.), *Interpretability Issues in Fuzzy Modeling*, Springer, Berlin, 2003.

[12] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena (eds.), *Accuracy Improvements in Linguistic Fuzzy Modeling*, Springer, Berlin, 2003.

[13] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms," *Fuzzy Sets and Systems*, vol. 65, no. 2/3, pp. 237-253, August 1994.

[14] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Trans. on Fuzzy Systems*, vol. 3, no. 3, pp. 260-270, August 1995.

[15] H. Ishibuchi, T. Murata, and I. B. Turksen, "Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems," *Fuzzy Sets and Systems*, vol. 89, no. 2, pp. 135-150, July 1997.

[16] H. Ishibuchi and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," *Fuzzy Sets and Systems*, vol. 141, no. 1, pp. 59-88, January 2004.

[17] H. Ishibuchi, T. Nakashima, and T. Murata, "Performance evaluation of fuzzy classifier systems for multi-dimensional pattern classification problems," *IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 29, no. 5, pp. 601-618, October 1999.

[18] H. Ishibuchi and T. Nakashima, "Improving the performance of fuzzy classifier systems for pattern classification problems with continuous attributes," *IEEE Trans. on Industrial Electronics*, vol. 46, no. 6, pp. 1057-1068, December 1999.

[19] H. Ishibuchi, T. Nakashima, and T. Murata, "Three-objective genetics-based machine learning for linguistic rule extraction," *Information Sciences*, vol. 136, no. 1-4, pp. 109-133, August 2001.

[20] H. Ishibuchi, T. Yamamoto, and T. Nakashima, "Hybridization of fuzzy GBML approaches for pattern classification problems," *IEEE Trans. on Systems, Man, and Cybernetics- Part B: Cybernetics*, vol. 35, no. 2, pp. 359-365, April 2005.

[21] H. Ishibuchi and Y. Nojima, "Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning," *International Journal of Approximate Reasoning* (in press).

[22] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*, Springer, Berlin, November 2004.

[23] H. Ishibuchi and T. Nakashima, "Effect of rule weights in fuzzy rule-based classification systems," *IEEE Trans. on Fuzzy Systems*, vol. 9, no. 4, pp. 506-515, August 2001.

[24] H. Ishibuchi and T. Yamamoto, "Rule weight specification in fuzzy rule-based classification systems," *IEEE Trans. on Fuzzy Systems*, vol. 13, no. 4, pp. 428-435, August 2005.

[25] O. Cordon, M. J. del Jesus, and F. Herrera, "A proposal on reasoning methods in fuzzy rule-based classification systems," *International Journal of Approximate Reasoning*, vol. 20, no. 1, pp. 21-45, January 1999.

[26] H. Ishibuchi, T. Nakashima, and T. Morisawa, "Voting in fuzzy rule-based systems for pattern classification problems," *Fuzzy Sets and Systems*, vol. 103, no. 2, pp. 223-238, April 1999.

[27] H. Ishibuchi, K. Nozaki, and H. Tanaka, "Distributed representation of fuzzy rules and its application to pattern classification," *Fuzzy Sets and Systems*, vol. 52, no. 1, pp. 21-32, November 1992.

[28] A. Gonzalez and R. Perez, "SLAVE: A genetic learning system based on an iterative approach," *IEEE Trans. on Fuzzy Systems*, vol. 7, no. 2, pp. 176-191, April 1999.

[29] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo, "Fast discovery of association rules," in U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park, pp. 307-328, 1996.

[30] F. Coenen, P. Leng, and L. Zhang, "Threshold tuning for improved classification association rule mining," *Lecture Notes in Computer Science 3518: Advances in Knowledge Discovery And Data Mining - PAKDD 2005*, pp. 216-225, Springer, Berlin, May 2005.

[31] F. Coenen and P. Leng, "Obtaining best parameter values for accurate classification," *Proc. of the 5th IEEE International Conference on Data Mining*, pp. 549-552, Houston, Texas, November 27-30, 2005.